



TITLE:

# On Chomsky Hierarchy of Palindromic Languages (Algebraic Systems and Theoretical Computer Science)

AUTHOR(S):

Domosi, Pal; Fazekas, Szilard; Ito, Masami

---

CITATION:

Domosi, Pal ...[et al]. On Chomsky Hierarchy of Palindromic Languages (Algebraic Systems and Theoretical Computer Science). 数理解析研究所講究録 2012, 1809: 1-6

ISSUE DATE:

2012-09

URL:

<http://hdl.handle.net/2433/194479>

RIGHT:

# On Chomsky Hierarchy of Palindromic Languages

*Pál Dömösi, \*Szilárd Fazekas<sup>†</sup>, Masami Ito<sup>‡,§</sup>*

*Abstract:* The characterization of the structure of palindromic regular and palindromic context-free languages is described by S. Horváth, J. Karhumäki, and J. Kleijn [5]. In this paper alternative proofs are given for these characterizations. Moreover, a simple observation is also given for palindromic context-sensitive (phrase-structural) languages.

## 1 Introduction

Characterization of palindromic regular and context-free languages is given by [5]. In this paper we give alternative proofs of these characterizations, moreover, we characterize the palindromic context-sensitive languages. (The palindromic phrase-structural languages have a trivial characterization).

## 2 Preliminaries

A *word* (over  $\Sigma$ ) is a finite sequence of elements of some finite non-empty set  $\Sigma$ . We call the set  $\Sigma$  an *alphabet*, the elements of  $\Sigma$  *letters*. If  $u$  and  $v$

---

\*College of Nyíregyháza, Institute of Mathematics and Informatics,  
H-4400 Nyíregyháza, Sóstói út 31/B, Hungary, e-mail: domosi@nyf.hu

<sup>†</sup>College of Nyíregyháza, Institute of Mathematics and Informatics,  
H-4400 Nyíregyháza, Sóstói út 31/B, Hungary, e-mail: domosi@nyf.hu

<sup>‡</sup>Kyoto Sangyo University, Faculty of Science, Kyoto 603-555, Japan,  
e-mail : ito@cc.kyoto-su.ac.jp

<sup>§</sup>The first two authors are grateful to JSPS (Japanese Society for Promotion of Science), Kyoto Sangyo University, and Nyíregyháza College for their constant support. The first author was also supported by Czech Ministry of Education, Youth and Sport, and Hungarian National Development Agency No CZ-1/2009.

are words over an alphabet  $\Sigma$ , then their *catenation*  $uv$  is also a word over  $\Sigma$ . Especially, for every word  $u$  over  $\Sigma$ ,  $u\lambda = \lambda u = u$ , where  $\lambda$  denotes the *empty word*. Two words  $u, v$  are said to be *conjugates* if there exists a word  $w$  with  $uw = wv$ . A nonempty word is called *primitive* if it is not a power of another word. Otherwise we speak about *nonprimitive word*. Thus  $\lambda$  is also considered as a nonprimitive word.

The *length*  $|w|$  of a word  $w$  is the number of letters in  $w$ , where each letter is counted as many times as it occurs. Thus  $|\lambda| = 0$ . By the *free monoid*  $\Sigma^*$  *generated by*  $\Sigma$  we mean the set of all words (including the *empty word*  $\lambda$ ) having catenation as multiplication. We set  $\Sigma^+ = \Sigma^* \setminus \{\lambda\}$ , where the subsemigroup  $\Sigma^+$  of  $\Sigma^*$  is said to be the *free semigroup generated by*  $\Sigma$ . Subsets of  $\Sigma^*$  are referred to as *languages* over  $\Sigma$ . Denote by  $|H|$  the *cardinality* of  $H$  for every set  $H$ . A language  $L$  is said to be *slender* if there exists a nonnegative integer  $c$  having  $|\{w \in L : |w| = n\}| \leq c$ .

For a nonempty word  $w = x_1 \cdots x_n$ , where  $x_1, \dots, x_n \in \Sigma$ , we denote its *reverse*,  $x_n \cdots x_1$ , by  $w^R$ . Moreover, by definition, let  $\lambda = \lambda^R$ , where  $\lambda$  denotes the empty word of  $\Sigma^*$ . We say that a word  $w$  is a *palindrome* (or *palindromic*) if  $w = w^R$ . Further, we call a language  $L \subseteq \Sigma^*$  *palindromic* if all of its elements are palindromes.

A language  $L \subseteq \Sigma^*$  is called a *paired loop language* if it is of the form  $L = \{uv^nwx^ny|n \geq 0\}$  for some words  $u, v, w, x, y \in \Sigma^*$ .

Finally, as usual, we write a *generative grammar*  $G$  into the form  $G = (V, \Sigma, S, P)$ , where  $V$  and  $\Sigma$  are nonempty finite distinct sets, the *set of nonterminals*, and the *set of terminals*,  $S \in V$  is the *start symbol*, and  $P$  is the finite set of *derivation rules*. For every *sentential form*  $W \in (V \cup \Sigma)^*$ ,  $L_G(W)$  denotes the *language generated by*  $W$ , and  $L(G) (= L_G(S))$  denotes the language *generated by*  $G$ .

We shall use the following classical results.

**Theorem 1** [1] *Let  $L$  be a regular language. Then there is a constant  $n$  such that if  $z$  is any word in  $L$ , and  $|z| \geq n$ , we may write  $z = uvw$  in such a way that  $|uv| \leq n$ ,  $|v| \geq 1$ , and for all  $i \geq 0$ ,  $uv^i w$  is in  $L$ . Furthermore,  $n$  is no greater than the number of states of the finite automaton with minimal states accepting  $L$ .*  $\square$

**Theorem 2** *The family of context-free languages is closed under the inverse homomorphism.*  $\square$

**Theorem 3** [1] *The language  $L \subseteq \Sigma^*$  is context-free if and only if for every regular language  $R \subseteq \Sigma^*$ ,  $L \cap R$  is context-free.*  $\square$

**Theorem 4** [4] *Given an alphabet  $\Sigma$ , a nonempty word  $w \in \Sigma^+$ , each context-free language  $L \subseteq w^*$  is regular having the form*

$$\cup_{i=1}^k w^{m_i} (w^{n_i})^* \text{ for some } m_1, n_1, \dots, m_k, n_k \geq 0. \quad (1)$$

$\square$

**Theorem 5** [6, 7, 9] *Every slender context-free language is a finite disjoint union of paired loop languages.*  $\square$

The following statement is well-known.

**Proposition 6** *Given a context-free grammar  $G = (V, \Sigma, S, P)$ , a sentential form  $W \in (V \cup \Sigma)^*$ , the language  $S_G(W)$  is also context-free.*  $\square$

**Theorem 7** [10] *Given a positive integer  $i$ , a pair  $u, v \in \Sigma^+$ , let  $uv = p^i$  for some primitive word  $p \in \Sigma^+$ . Then  $vu = q^i$  for a primitive word  $q$ .*  $\square$

**Theorem 8** [8] *If  $uv = vq$ ,  $u \in \Sigma^+$ ,  $v, q \in \Sigma^*$ , then  $u = wz$ ,  $v = (wz)^k w$ ,  $q = zw$  for some  $w \in \Sigma^*$ ,  $z \in \Sigma^+$  and  $k \geq 0$ .*  $\square$

**Theorem 9** [8] *The words  $u, v \in \Sigma^*$  are conjugates if and only if there are words  $p, q \in \Sigma^*$  with  $u = pq$  and  $v = qp$ .*  $\square$

**Theorem 10** [2] *Let  $u, v \in \Sigma^*$ .  $u, v \in w^+$  for some  $w \in \Sigma^+$  if and only if there are  $i, j \geq 0$  so that  $u^i$  and  $v^j$  have a common prefix (suffix) of length  $|u| + |v| - \gcd(|u|, |v|)$ .*  $\square$

We shall use the following direct consequence of this result.

**Theorem 11** *If two non-empty words  $p^i$  and  $q^j$  share a prefix of length  $|p| + |q|$ , then there exists a word  $r$  such that  $p, q \in r^+$ .*  $\square$

### 3 Results

We start with alternative proofs of some results of S. Horváth, J. Karhumäki, J. Kleijn [5].

First we turn to consider regular languages.

**Theorem 12** [5] *A regular language  $L \subseteq \Sigma^*$  is palindromic if and only if it is a union of finitely many languages of the form*

$$L_p = \{p\}, L_{q,r,s} = qr(sr)^*q^R, (p, q, r, s \in \Sigma^*), \quad (2)$$

where  $p, r$  and  $s$  are palindromes.

*Proof:* Clearly, any finite union of languages in (2) is both palindromic and regular. Conversely, let  $L$  be a palindromic regular language and  $n$  be the language-specific constant from Theorem 1. Naturally, there are finitely many words shorter than  $n$ , those will form the languages  $L_p$ . For any suitably long word  $w \in L$ , according to Theorem 1, we have a factorization  $w = qvz$ , with  $0 < |qv| \leq n$  and  $v \neq \lambda$ , such that  $qv^iz \in L$ , for any  $i \geq 0$ . The two cases being symmetric, we may assume  $|q| \leq |z|$ , i.e.,  $z = xq^R$ , for some  $x \in \Sigma^*$ , with  $v^ix$  being a palindrome. This gives us  $x = r(v^R)^j$ , for some  $r$  with  $v^R = sr$  and some  $j \geq 0$ . But, for large enough  $i$ ,  $v^ix$  ends in  $sx = (v^Rv^R)^R x = (r^R s^R)^2 r (v^R)^j$  and it starts with  $v^{j+2}$ , so we instantly get  $v = r^R s$  and thus  $s = s^R$ . It also follows, that  $v^R = s^R r$  and  $v^R = s^R r^R$ , hence  $r$  is a palindrome, too. Then, our original word  $w$  can be written as  $qr(sr)^{j+k}q^R$ . A similar decomposition, according to Theorem 1 is bound to exist for all words longer than  $n$ . All parts of the decomposition,  $q, r$  and  $s$  are shorter than  $n$ , therefore there are finitely many triplets like this.  $\square$

Next we prove the following simple observation.

**Proposition 13** *Given a pair of positive integers  $i, j$ , let  $p, r, u, w \in \Sigma^*, v \in \Sigma^+$  be arbitrary with  $|p| \leq |u|, |r| \leq |w|$  and let  $q \in \Sigma^+$  be a primitive word having  $|v^j| \geq |v| + 3|q|$  such that  $pq^i r = uv^j w$ . Then there exists a positive integer  $k$  such that  $v$  and  $q^k$  conjugate.*

*Proof:* By our assumptions, there exists a pair of factorizations  $u = pu', v = v'q$  such that  $q^i = u'v^jv'$ . Because  $|v^j| \geq |v| + 3|q|$ ,  $|u'v'| = |q^i| - |v^j| \leq$

$|q^i| - |v| - 3|q| < |q^{i-3}|$ , there are a positive integer  $n$ , a suffix  $q_2$  and a prefix  $q_3$  of  $q$  such that  $v^j = q_2 q^n q_3$ . Hence  $v^j = q_2 (q_1 q_2)^n q_3 = (q_2 q_1)^n q_2 q_3$  for some decomposition  $q = q_1 q_2$  and prefix  $q_3$  of  $q$ . By our conditions,  $|v^j| - |q_3| \geq |v| + 3|q| - |q_3| \geq |v| + 2|q| > |v| + |q|$ . Therefore, applying Theorem 11, we obtain  $v, q_2 q_1 \in z^+$  for some primitive word  $z \in \Sigma^+$ . By Theorem 7,  $q_2 q_1$  is also primitive. Therefore,  $z = q_2 q_1$ . Hence  $v = (q_2 q_1)^k$  for some  $k > 0$ . Then Theorem 9 implies that  $v$  and  $q^k$  conjugate.  $\square$

Now we continue with palindromic context-free languages.

**Theorem 14** [5] *Every palindromic context-free language is linear.*

*Proof:* Let  $G = (V, \Sigma, S, P)$  be a context-free grammar generating the palindromic language  $L$ . Without loss of generality we can assume that  $V$  is reduced, i.e., for every  $X \in V$ ,  $L_G(X) \neq \emptyset$ . In particular, we may assume for every  $X \in V$ ,  $|L_G(X)| = \infty$ . Indeed, if  $|L_G(X)| < \infty$ , then we can eliminate the derivation rules

$$Y \rightarrow W_1 X W_2 X \cdots W_n X W_{n+1}, X \rightarrow W \in P,$$

$W, W_1, W_2, \dots, W_{n+1} \in ((V \setminus \{X\}) \cup \Sigma)^*$  by new derivation rules of the form

$$Y \rightarrow W_1 w_1 W_2 w_2 \cdots w_n W_{n+1}, w_1, \dots, w_n \in L_G(X).$$

It can also be assumed that for every  $X \rightarrow W \in P$ , there are at most two (not necessarily different) nonterminals appearing in  $W$ . Indeed, if  $X \rightarrow u_1 A_1 \cdots u_n A_n u_{n+1} \in P$  with  $X, A_1, \dots, A_n \in V, u_1, \dots, u_n \in \Sigma^*, n > 2$  then we can eliminate this derivation rule by the following new derivation rules using some new nonterminals  $A'_1, \dots, A'_{n-1}$ :

$$X \rightarrow u_1 A_1 u_2 A'_2, A'_2 \rightarrow A_2 u_3 A'_3, \dots, A'_{n-2} \rightarrow A_{n-2} u_{n-1} A'_{n-1}, A'_{n-1} \rightarrow A_{n-1} u_n.$$

Next we show that the derivation rules of the form  $X \rightarrow p A q B r$  with  $p, q, r \in \Sigma^*, A, B \in V$  can be eliminated.

First we establish that for every  $q_1, q_2 \in \Sigma^*, A \xrightarrow{*}_G q_1, A \xrightarrow{*}_G q_2, q_1 \neq q_2$  implies  $|q_1| \neq |q_2|$ . Similarly, for every  $r_1, r_2 \in \Sigma^*, B \xrightarrow{*}_G r_1, B \xrightarrow{*}_G r_2, r_1 \neq r_2$  implies  $|r_1| \neq |r_2|$ . Because  $G$  is reduced, there are  $u, y \in \Sigma^*$  having  $S \xrightarrow{*}_G u X y$ . Therefore,  $A \xrightarrow{*}_G q_1$  and  $A \xrightarrow{*}_G q_2$  imply that for every  $r' \in L_G(B)$ ,

$upq_1qr'ry, upq_2qr'ry \in L$ , i.e., both of them are palindromes. This is impossible if  $|q_1| = |q_2|$  with  $q_1 \neq q_2$ . Similarly,  $B \xrightarrow{*}_G r_1$  and  $B \xrightarrow{*}_G r_2$  imply that for every  $q' \in L_G(A)$ ,  $upq'qr_1ry, upq'qr_2ry \in L$ , i.e., both of them are palindromes. This is impossible if  $|r_1| = |r_2|$  and  $r_1 \neq r_2$ .

This means that all of the languages  $L_G(A), L_G(B)$  are slender context-free languages. Using Theorem 5,  $X \rightarrow pAqBr$  can be eliminated by considering some new nonterminals  $A_1, \dots, A_m, B_1, \dots, B_n$  and for every  $i = 1, \dots, m, j = 1, \dots, n$ , new derivation rules  $X \rightarrow pu_iA_iy_iqu_jB_jy_jr$ ,  $A_i \rightarrow v_iA_ix_i, A_i \rightarrow w_i, B_j \rightarrow v'_jB_jx'_j, B \rightarrow w' \in P$ , where  $u_i, v_i, w_i, x_i, y_i, u'_i, v'_i, w'_i, x'_i, y'_i \in \Sigma^*$ . Therefore, we may suppose that for every  $X \rightarrow pAqBr \in P, A, B \in V, p, q, r \in \Sigma^*, A \rightarrow vAx, A \rightarrow w, B \rightarrow v'Ax', B \rightarrow w' \in P, v, w, x, v', w', x' \in \Sigma^*$  and  $A, B$  do not appear on the left side of any other derivation rules. Thus

$$L_G(pAqBr) = \{pv^iwx^iqv'^jw'x'^jr \mid i, j \geq 0\}.$$

To our statement it is enough to prove that at least one of the following equalities is true:  $wx = xw, v'w' = w'v'$ . Indeed, if  $wx = xw$  then  $X \rightarrow pAqBr \in P$  can be eliminated by linear derivation rules as follows:

omit the derivation rules  $X \rightarrow pAqBr, A \rightarrow vAx, A \rightarrow w$  and let  $X \rightarrow pC, C \rightarrow vxC, C \rightarrow wqB$  be new ones with the new nonterminal  $C$ ;

similarly, if  $v'w' = w'v'$  then  $X \rightarrow pAqBr \in P$  can be eliminated by the following linear derivation rules:

omit the derivation rules  $X \rightarrow pAqBr, B \rightarrow v'Bx', B \rightarrow w'$  and let  $X \rightarrow pCr, C \rightarrow Cv'x', C \rightarrow Aqw'$  be new ones with the new nonterminal  $C$ .

Therefore, if one of  $v, x, v', x'$  is empty then we are ready. Thus assume that none of  $v, x, v', x'$  is empty.

Let  $S \xrightarrow{*}_G u'Xy'$  for some  $u', y' \in \Sigma^*$  with  $|u'| \geq |y'|$ . Then for every  $z \in L_G(pAqBr)$ ,  $u'zy' \in L$ . Hence  $u'zy' = (u'zy')^R$ . Therefore,  $u' = y'^Ru$  for some  $u \in \Sigma^*$  such that for every  $z \in L_G(pAqBr)$ ,  $uz = (uz)^R$ .

Recall that  $L_G(pAqBr) = \{pv^iwx^iqv'^jw'x'^jr \mid i, j \geq 0\}$  with  $|x'| > 0$ .

Let  $z$  denote the primitive root of  $v$ . Moreover, let  $k$  be a nonnegative integer such that  $|up| \leq |x'^k r|$ . First choose  $i$  and  $j$  such that  $|x'^{k+1}| + 3|z| \leq |x'^{j-1}| \leq |upv^i| < |x'^j r|$ . Hence  $|x'| + 3|z| \leq |x'^{j-k-1}|$  and  $(upv^i)^R = x'_2 x'^R r$  for

some suffix  $x'_2$  of  $x'$ . Recall that  $|up| \leq |x'^k r|$ . Therefore, applying Proposition 13,  $x'$  and a power of  $z^R$  conjugate.

Now we choose  $i$  and  $j$  such that  $|v'w'x'^{j-\ell}r| + 3|z| \leq |v'^{j-\ell}w'x'^{j-\ell}r| \leq |upv^i| \leq |x'^k v^i| < |v'^j w'x'^j r|$ . Hence  $|v'| + 3|z| \leq |v'^{j-\ell}|$  and  $(upv^i)^R = v = 2'v'^{j-1}w'x'^{j-1}r$ , where  $v'_2$  is a suffix of  $v'$ . Applying Proposition 13 again, we obtain that  $v'$  and a power of  $z^R$  also conjugate.

Consider a pair of nonnegative integers  $s, t \geq 0$  such that  $|upv^s| = |x'_2 x'^t r|$  for some suffix  $x'_2$  of  $x'$ . Then, by our assumptions, for every  $i \geq s, j \geq t+1$ ,  $v^{i-s}w'x'^i qv'^j w'x'^{j-t-1}x'_1$  with  $x' = x'_1 x'_2$  is a palindrome. Consider a positive integer  $j$  such that  $|v| < |v'^j|$  and let  $i$  be given such that  $i-s$  is the smallest positive integer having  $|w'x'^{j-t-1}x'_1| \leq |v^{i-s}|$ . Obviously, then  $|v^{i-s}| \leq |v'^j w'x'^{j-t-1}x'_1|$ . Thus we may assume  $(v^{i-s})^R = v'_2 v'^{j-\ell-1}w'x'^{j-t-1}x'_1$  for some  $\ell \geq 0$ , for for some suffix  $v'_2$  of  $v'$  and some prefix  $x'_1$  of  $x'$ . Recall that  $v'$  and a power of  $z^R$ , moreover,  $x'$  and a power of  $z^R$  conjugate. Hence  $w' = z_1^R z^a z_2^R$  for some nonnegative  $a$ , a suffix  $z_2$  and a prefix  $z_1$  of the primitive root  $z$  of  $v$ . Moreover, because  $v'$  and a power of  $z^R$ , and also  $x'$  and a power of  $z^R$  conjugate,  $(v^{i-s})^R = v'_2 v'^{j-\ell-1}w'x'^{j-t-1}x'_1$  and  $w' = z_1^R (z^R)^a z_2^R$  imply  $v' = (z_1^R z_4^R)^b$  and  $x' = (z_3^R z_2^R)^c$  for some  $b, c > 0$  such that  $z_4^R z_1^R = z_2^R z_3^R = z^R$ . Hence

$$\begin{aligned} upv^i w'x'^i qv'^j w'x'^j r &= upv^i w'x'^i p(r^R (z_2 z_3)^{bj} z_2^a z_1 (z_4 z_1)^{cj})^R = \\ upv^i w'x'^i q(r z_2 (z^{bj+a+cj} z_1)^R) &= upv^i w'x'^i q z_1^R (z^R)^{bj+a+cj} z_2^R r = \\ upv^i w'x'^i q z_1^R ((z^R)^b)^j ((z^R)^c)^j (z^R)^a z_2^R r. \end{aligned}$$

Choose  $\bar{q} = q z_1, \bar{v}' = (z^R)^b, \bar{w}' = (z^R)^a, \bar{x}' = (z^R)^c, \bar{r} = z_2^R r$ .

Modify the grammar  $G$  such that omit the derivation rules  $X \rightarrow pAqBr, B \rightarrow v'Bx', B \rightarrow w'$  and let  $X \rightarrow pA\bar{q}C\bar{r}, C \rightarrow \bar{v}'C\bar{x}', C \rightarrow \bar{w}'$  be new derivation rules with the new nonterminal  $C$ . Obviously,  $L_G(pAqBr) = L_{G'}(pA\bar{q}C\bar{r})$ , and thus,  $L(G) = L(G')$ . On the other hand, by our constructions, for every  $i, j \geq 0, \bar{v}'\bar{w}' = \bar{w}'\bar{v}'$ . Therefore, as we have already seen, the derivation rules having the form  $X \rightarrow pA\bar{q}B\bar{r}, B \rightarrow \bar{v}'B\bar{x}', B \rightarrow \bar{w}'$  can be eliminated by the following new ones  $X \rightarrow pC\bar{r}, C \rightarrow C\bar{v}'\bar{x}', C \rightarrow Aq\bar{w}'$ , where  $C$  is a new nonterminal.

We assumed in the proof that  $S \xrightarrow{*}_G u'Xy'$  such that  $|u'| \geq |v'|$ . Changing the roles of the right and left sides of the discussed strings, we can also eliminate the derivation rules of the form  $X \rightarrow pAqBR$  if  $S \xrightarrow{*}_G u'Xy'$  for some  $u', v' \in \Sigma^*$  with  $|y'| \geq |u'|$ . Thus we receive that  $L(G)$  can be generated by a linear grammar.  $\square$

**Lemma 15** *Given an alphabet  $\Sigma$ , words  $v, z \in \Sigma^*$ , a non-empty word  $w \in$*



$\Sigma^+$ , each context-free language  $L \subseteq vw^*z$  is regular having the form

$$v(\cup_{i=1}^k w^{m_i}(w^{n_i})^*)z \text{ for some } m_1, n_1, \dots, m_k, n_k \geq 0. \quad (3)$$

*Proof:* Let  $a, b, c$  distinct symbols and consider a homomorphism  $\psi : \{a, b, c\} \rightarrow \Sigma^*$  with  $\psi(a) = v, \psi(b) = w, \psi(c) = z$ . Then  $\psi^{-1}(L) \cap ab^*c = \{ab^k c \mid vw^k z \in L, k \geq 0\}$ . On the other hand, using that  $ab^*c$  is obviously a regular language, Theorem 2 and Theorem 3 imply that  $\psi^{-1}(L) \cap ab^*c$  is also context-free. Let  $\psi' : \{a, b, c\} \rightarrow b^*$  be a homomorphism with  $\psi'(a) = \psi'(c) = \lambda$  and  $\psi'(b) = b$ . By Corollary 2,  $\psi'(\psi^{-1}(L) \cap ab^*c)$  is also context-free. On the other hand,  $\psi'(\psi^{-1}(L) \cap ab^*c) = \{b^k \mid vw^k z \in L, k \geq 0\}$ , therefore, by Theorem 4, it is regular which can be written into the form  $\cup_{i=1}^k b^{m_i}(b^{n_i})^*$  for some  $m_1, n_1, \dots, m_k, n_k \geq 0$ . This fact and the equality  $\psi'(\psi^{-1}(L) \cap ab^*c) = \{w^k \mid vw^k z \in L, k \geq 0\}$  implies that  $L$  is regular having the form as in (3).  $\square$

**Lemma 16** *Every palindromic context-free language can be generated by a grammar  $G = (V, \Sigma, S, P)$  having  $P \subseteq \{X \rightarrow aYa \mid X, Y \in V, a \in \Sigma\} \cup \{X \rightarrow a \mid X \in V, a \in \Sigma\} \cup \{X \rightarrow \Sigma\}$ .*

*Proof:* Consider an arbitrary palindromic context-free language  $L$ . By Theorem 14, we have that  $L$  is linear. Thus there exists a linear grammar  $G = (V, \Sigma, S, P)$ . Without restriction, we may assume that  $G$  is reduced, moreover,  $P \subseteq \{X \rightarrow aYb \mid X \in V, Y \in V \cup \{\lambda\}, a, b \in \Sigma \cup \{\lambda\}, ab \neq \lambda\}$ . Indeed, if  $X \rightarrow paYbq \in P$  with  $p, q \in \Sigma^*, pq \in \Sigma^+, a, b \in \Sigma \cup \{\lambda\}, ab \neq \lambda, Y \in V \cup \{\lambda\}$ , then we can eliminate the derivation rule  $X \rightarrow paYbq \in P$  by introducing a new nonterminal symbol  $Z$  and the new derivation rules  $X \rightarrow pZq, Z \rightarrow aYb$ . Thus we get in finite-many steps that all derivation rules have the form  $X \rightarrow aYb, X \in V, a, b \in \Sigma \cup \{\lambda\}, Y \in V \cup \{\lambda\}$ .

Clearly, then

$$L = \cup \{ \{p\} L_G(X) \{q\} \mid S \xrightarrow{*}_G pXq, X \in V, p, q \in \Sigma^*, |p|, |q| \leq |V| \}. \quad (4)$$

Next we prove that all of the derivation rules having one of the forms  $X \rightarrow aY, X, Y \in V, a \in \Sigma$  or  $X \rightarrow aY, X, Y \in V, a \in \Sigma$  can be eliminated.

We say that a nonterminal  $X \in V$  is *non-balanced* if there are  $p, q \in \Sigma^*$  with  $|p| \neq |q|$  such that  $X \xrightarrow{*}_G pXq$ . Otherwise, we say that  $X$  is *balanced*. Now we eliminate the non-balanced nonterminals. Consider a non-balanced

nonterminal  $X$ , as above. Let us assume  $X$  appears in a derivation at some point as  $S \Rightarrow uXv$ . Then because  $X \Rightarrow pXq$ , we get  $S \Rightarrow up^iXq^iv$ , for all  $i \geq 1$ . Without loss of generality, we may assume  $|u| \leq |v|$ , that is, since the derived word will be a palindrome,  $v = wu^R$ , for some  $w \in \Sigma^*$ . Now, to keep arguments simple, let  $X$  stand for any word in  $L_G(X)$ . So, we know that  $p^iXq^iw$  is a palindrome for any positive  $i$ . For large enough  $i$ , this gives us that  $w^R = p^jp_1$ , for some  $j \geq 0$  and  $p_1 \in \Sigma^*$  prefix of  $p$ , hence  $p^iXq^ip_1^R(p^R)^j$  is a palindrome. Again, if  $i$  was big enough for  $|p^i| > |q^2p_1^R(p^R)^j|$ , then by Theorem 10, we get that for a decomposition  $q_1q_2$  of  $q^R$ , its conjugate  $q_2q_1$  has the same primitive root as  $p$ , i.e., there exists some primitive word  $z \in \Sigma^+$ ,  $m, n \geq 1$ , such that  $q_2q_1 = z^m$  and  $p = z^n$ . Rewriting  $p^iXq^ip_1^R(p^R)^j$  with these powers of  $z$ , we have  $z^{ni}X(q_2^Rq_1^R)^ip_1(z^R)^{nj} = z^{ni}Xq_2^R(q_1^Rq_2^R)^{i-1}q_1^Rp_1(z^R)^{nj} = z^{ni}Xq_2^R(z^R)^{m(i-1)}q_1^Rp_1(z^R)^{nj}$  is a palindrome, therefore  $z^{n(i-j)}Xq_2^R(z^R)^{m(i-1)}q_1^Rp_1$  is, as well. This means  $p_1^Rq_1z^2$  is a prefix of  $z^{n(i-j)}$ , and we can apply Theorem 10 again to get that, since  $z$  is primitive,  $p_1^Rq_1 = z^k$ , for some integer  $k$ . Since  $p_1^R$  is a suffix of  $p^R = (z^R)^n$  and  $q_1$  is a suffix of  $z^m$ , there exist non-negative integers  $i_1, i_2$  and  $z'_r$  suffix of  $z^R$ ,  $z'$  suffix of  $z$ , such that  $z'_r(z^R)^{i_1}z'z^{i_2} = z^k$ . From here, there is some prefix  $z''_r$  of  $z^R$ , with  $z''_rz'_r = z^R$ ,  $z'_rz''_r = z$ , so both  $z''_r$  and  $z'_r$  are palindromes and so are  $p_1 = z'_r(z''_rz'_r)^{i_1}$  and  $q_1 = (z''_rz'_r)^{k-i_1-1}z''_r$ . But  $q_2q_1 = z^m = (z'_rz''_r)^m$ , so  $q_2 = z'_r(z''_rz'_r)^{m-k+i_1+1}$ . From here,  $z^{ni}X(q_2^Rq_1^R)^ip_1(z^R)^{nj} = (z'_rz''_r)^{ni}X(z'_rz''_r)^{mi}z'_r(z''_rz'_r)^{i_1}(z''_rz'_r)^{nj} = (z'_rz''_r)^{ni}X(z'_rz''_r)^{mi+i_1+nj}z'_r$  is a palindrome for all  $i \geq 1$ . As our original assumption was  $|p| \neq |q|$ , i.e.,  $m \neq n$ , for a large enough  $i$ , the word  $X$  will be entirely to the left or right from the center of a palindrome of the form  $(z'_rz''_r)^{j_1}X(z'_rz''_r)^{j_2}z'_r$ . Since  $z'_rz''_r$  is primitive, the center of the palindrome has to be exactly  $z'_r$  or  $z''_r$ , and this means that  $X \in (z'_rz''_r)^+$ . Then, the language  $L_G(X)$  is isomorphic to a unary context-free language, hence it is regular with rules of the form  $X \rightarrow (z'_rz''_r)^{m+n}X$ . This way, in our original grammar we can replace all rules with  $X$  on the left with balanced rules  $X \rightarrow (z'_rz''_r)^{\frac{m+n}{2}}X(z'_rz''_r)^{\frac{m+n}{2}}$ , or if  $m+n$  is odd, with rules  $X \rightarrow (z'_rz''_r)^{m+n}X(z'_rz''_r)^{m+n}$  and  $X \rightarrow (z'_rz''_r)^{m+n}|\lambda$ , etc.  $\square$

**Lemma 17** *Every palindromic context-free language can be generated by a grammar  $G = (V, \Sigma, S, P)$  having  $P \subseteq \{X \rightarrow aYa \mid X, Y \in V, a \in \Sigma\} \cup \{X \rightarrow a \mid X \in V, a \in \Sigma\} \cup \{X \rightarrow \Sigma\}$ .*

*Proof:* Consider an arbitrary palindromic context-free language  $L$ . By Theorem 14, we have that  $L$  is linear. Thus there exists a linear grammar

$G = (V, \Sigma, S, P)$ . Without restriction, we may assume that  $G$  is reduced, moreover,  $P \subseteq \{X \rightarrow aYb \mid X \in V, Y \in V \cup \{\lambda\}, a, b \in \Sigma \cup \{\lambda\}, ab \neq \lambda\}$ . Indeed, if  $X \rightarrow paYbq \in P$  with  $p, q \in \Sigma^*, pq \in \Sigma^+, a, b \in \Sigma \cup \{\lambda\}, ab \neq \lambda, Y \in V \cup \{\lambda\}$ , then we can eliminate the derivation rule  $X \rightarrow paYbq \in P$  by introducing a new nonterminal symbol  $Z$  and the new derivation rules  $X \rightarrow pZq, Z \rightarrow aYb$ . Thus we get in finite-many steps that all derivation rules have the form  $X \rightarrow aYb, X \in V, a, b \in \Sigma \cup \{\lambda\}, Y \in V \cup \{\lambda\}$ .

Clearly, then

$$L = \cup\{\{p\}L_G(X)\{q\} \mid S \xrightarrow{*}_G pXq, X \in V, p, q \in \Sigma^*, |p|, |q| \leq |V|\}. \quad (5)$$

Next we prove that all of the derivation rules having one of the forms  $X \rightarrow aY, X, Y \in V, a \in \Sigma$  or  $X \rightarrow aY, X, Y \in V, a \in \Sigma$  can be eliminated.

We say that a nonterminal  $X \in V$  is *non-balanced* if there are  $p, q \in \Sigma^*$  with  $|p| \neq |q|$  such that  $X \xrightarrow{*}_G pXq$ . Otherwise, we say that  $X$  is *balanced*. Now we eliminate the non-balanced nonterminals. To complete our proof, for every  $X \in V$  with  $X \rightarrow vXx$  and  $|v| \neq |x|$ , first we eliminate the productions having the form  $X \rightarrow aYb, Y \in V \cup \{\lambda\}, a, b \in \Sigma \cup \{\lambda\}$ .

Obviously, then,  $S \xrightarrow{*}_G pXq, X \xrightarrow{+}_G w, X \xrightarrow{+}_G vXx$  imply that for every  $i \geq 0$ ,  $pv^iwx^iq$  is a palindrome.

Therefore, for every non-negative integer  $m$ , there exists a pair  $k, \ell \geq m$  with  $pv^k = (x_2x^\ell q)^R$  for some suffix  $x_2$  of  $x$ . Indeed, if  $|pv^m| \geq |x^mq|$  then  $pv^m = (x_2x^\ell q)^R$  for some  $\ell \geq m$  and suffix  $x_2$  of  $x$ . Similarly, if  $|pv^m| < |x^mq|$  then  $pv^k = (x_2x^mq)^R$  for some  $k \geq m$  and suffix  $x_2$  of  $x$ .

Suppose  $|v| > |x|$ . Then there exists a non-negative integer  $i$  with  $|pv^i| \geq |wx^iq|$ . Hence,  $pv^{i-j-1}v_1 = (v_2v^jwx^iq)^R$  for some factorization  $v = v_1v_2$  and  $j \geq 0$ . But then  $v_2v_1 = (v_2v_1)^R$ , and thus,  $v_2(v_1v_2)^jwx^iq = (v_2v_1)^{i-j-1}v_1^R p^R$ , i.e.,  $w$  is a prefix of  $v_1(v_2v_1)^{i-j-1}v_1^R p^R$ . Hence,  $w = z_1z^kz_2$  for some  $k \geq 0$ , where  $z_1$  is a proper prefix of  $v_1v_2v_1$ ,  $z \in (v_2v_1)^*$ , and  $z_2$  is a proper prefix of  $v_2v_1v_1^R p^R$ .

Next we assume  $|v| < |x|$ . Then for an appropriate non-negative integer  $i$ ,  $pv^iwx^jx_1 = (x_2x^{i-j-1}q)^R$  for some factorization  $x = x_1x_2$  and non-negative integer  $j \geq 0$ . This implies  $x_2x_1 = (x_2x_1)^R$  and that  $pv^iwx^jx_1 = q^R x_2^R x_2^R (x_2x_1)^{i-j-2}x_1^R$ , i.e.,  $w$  is a suffix of  $q^R x_2^R (x_2x_1)^{i-2j-2}x_1^R$ .

Hence,  $w = z_1z^kz_2$  for some  $k \geq 0$ , where  $z_1$  is a proper suffix of  $q^R x_2^R x_2x_1$ ,  $z \in (x_2x_1)^*$ , and  $z_2$  is a proper suffix of  $x_2x_1x_1^R$ .

In both cases we receive that  $w \in z_1z^*z_2$  for an appropriate primitive palindrome  $z$  and words  $z_1, z_2 \in \Sigma^*$ .

By Proposition 6 and Lemma 15,

$$L_G(X) = z_1(\cup_{i=1}^k z^{m_i}(z^{n_i})z_2 \text{ for some } m_1, n_1, \dots, m_k, n_k \geq 0. \quad (6)$$

Introducing some new nonterminals and derivation rules, such that each  $Z$  of them has the property that  $Z \xrightarrow{*}_G pZq, P, q \in \Sigma^*$  implies  $|p| = |q|$ . we can derive the language  $L_G(X)$  as follows.

Omit all derivation rules of the form  $X \rightarrow w, w \in (V \cup \Sigma)^*$ , and let  $z_1 = a_1 \cdots a_k, z_2 = b_1 \cdots b_\ell, z = c_1 \cdots c_m$ . Consider the new derivation rules  $X \rightarrow a_1 X_1, X_1 \rightarrow a_2 X_2, \dots, X_{k-1} \rightarrow a_k X_k, X_k \rightarrow Y_\ell b_\ell, Y_\ell \rightarrow Y_{\ell-1} b_{\ell-1}, \dots, Y_2 \rightarrow Y_1 b_1$ , where  $X_1, \dots, X_k, Y_1, \dots, Y_\ell$  new nonterminals. Obviously, then  $X \xrightarrow{*}_G z_1 Y_1 z_2$ . Now, let  $m, n \geq 0$  with  $m + m > 0, z = c_1 \dots c_s d c_s \dots c_1, c_1, \dots, c_s \in \Sigma, d \in \Sigma \cup \{\lambda\}$ . We distinguish the following cases.

*Case 1* If  $m = 0$  then let  $Y_1 \rightarrow \lambda$ .

*Case 2* If  $m = 2i$  for some  $i > 0$ , then let  $Y_1 \rightarrow c_1 A_1 c_1, A_1 \rightarrow c_2 A_2 c_2, \dots, A_{s-2} \rightarrow c_{s-1} A_{s-1} c_{s-1}, A_{s-1} \rightarrow c_s A_s c_s, A_s \rightarrow dB_1 d, B_1 \rightarrow c_s A_{s+1} c_s, A_{s+1} \rightarrow c_{s-1} A_{s+2} c_{s-1}, \dots, A_{2s-1} \rightarrow c_1 A_{2s} c_1, A_{2s} \rightarrow dB_2 d, \dots, B_{2i-1} \rightarrow c_s A_{2(i-1)s+1} c_s, A_{2(i-1)s+1} \rightarrow c_{s-1} A_{2(i-1)s+2} c_{s-1}, \dots, A_{2is-1} \rightarrow c_1 A_{2is} c_1, A_{2is} \rightarrow \lambda$  be new derivation rules with some new nonterminals  $A_1, \dots, A_{2is}, B_1, \dots, B_{2i-1}$ .

*Case 3* If  $m = 2i + 1$  for some  $i > 0$ , then similarly as before, let  $Y_1 \rightarrow c_1 A_1 c_1, \dots, A_{s-1} \rightarrow c_s A_s c_s, A_s \rightarrow dB_1 d, \dots, B_{2i-1} \rightarrow c_s A_{2(i-1)s+1} c_s, \dots, A_{2is-1} \rightarrow c_1 A_{2is} c_1$ . Moreover, let  $A_{2is} \rightarrow c_1 A_{2is+1} c_1, A_{2is+1} \rightarrow c_2 A_{2is+2} c_2, \dots, A_{(2i+1)s-1} \rightarrow c_s A_{2(i+1)s} c_s, A_{2(i+1)s} \rightarrow d$  be new derivation rules containing some new nonterminals  $A_1, \dots, A_{2(i+1)s}, B_1, \dots, B_{2i-1}$ .

*Case 4* Finally, if  $m = 1$ , then analogously to the previous case, let  $Y_1 \rightarrow c_1 A_1 c_1, A_1 \rightarrow c_2 A_2 c_2, \dots, A_{s-1} \rightarrow c_s A_s c_s, A_s \rightarrow d$  be new derivation rules, where  $A_1, \dots, A_s$  be new nonterminals.

Obviously, in all of the above cases,  $X \xrightarrow{*}_G z_1 z^m z_2$ . Therefore, if  $n = 0$ , then  $z_1 z^m (z^n)^* z_2 \subseteq L_G(X)$ .

If  $n > 0$ , then we introduce a new derivation rule  $A_{ms} \rightarrow c_1 A'_1 c_1$ , moreover, analogously to the above Cases, distinguishing the cases  $n = 2j$  or  $n = 2j+1$  for some  $j > 0$ , or  $n = 1$ , we introduce further new derivation rules with some new nonterminals. In particular, for every above defined new derivation rule having one of the forms  $A_e \rightarrow c_f A_{e+1} c_f, A_{ns} \rightarrow \lambda, A_{ns} \rightarrow d, A_{gs} \rightarrow dB_g d, B_g \rightarrow c_s A_{2gs+1} c_s$ , we consider appropriate further new derivation rules with of the form, in order,  $A'_e \rightarrow c_f A'_{e+1} c_f$ ,

$A'_{ns} \rightarrow \lambda, A'_{ns} \rightarrow d, A'_{gs} \rightarrow dB'_gd, B'_g \rightarrow c_s A'_{2gs+1} c_s, A'_e \rightarrow c_f A'_g c_f$ , where  $A'_e, A'_{e+1}, A'_{ns}, A'_{gs}, A'_{2gs+1}, B'_g$  denote new nonterminals.

Finally, we also consider a new derivation rule  $A'_{ns} \rightarrow Y_1$  with the further new nonterminal  $A'_{ns}$ .

Obviously, in all cases,  $X \xrightarrow{*}_G z_1 z^m (z^n)^* z_2$ . Therefore,  $z_1 z^m (z^n)^* z_2 \subseteq L_G(X)$ . By (6),  $L_G(X)$  consists of finite-many languages having the above form. Therefore, we receive in finite-many steps that  $L_G(X)$  can be generated by new derivation rules containing only balanced nonterminal on their right-hand sides.

Now we assume that  $V$  contains only balanced nonterminals, i.e., for every derivation,  $X \xrightarrow{*}_G uXx, X \in V, u, x \in \Sigma^*, |u| = |x|$ . Then, for every  $X \in V, p, q \in \Sigma^*, S \xrightarrow{*}_G pXq$  implies  $||p| - |q|| < |V|$ . Indeed, assume the contrary and, for the simplicity, put  $X_0 = S$ . Then there exists a derivation

$$X_0 \xrightarrow{*}_G x_1 X_1 y_1 \xrightarrow{*}_G \cdots \xrightarrow{*}_G x_{n-1} X_{n-1} y_{n-1} \cdots y_1 \xrightarrow{*}_G x_1 \cdots x_n X_n y_n \cdots y_1, \quad (7)$$

where  $X_0, \dots, X_n \in V$ , and by our assumptions,  $x_1, \dots, x_n, y_1, \dots, y_n \in \Sigma \cup \{\lambda\}$ . On the other hand, if  $X_i = X_j$  for some  $i, j$  with  $1 \leq i < j \leq n$  then  $X_i \xrightarrow{*}_G x_{i+1} \cdots x_j X_i y_j \cdots y_{i+1}$  also holds.

If  $|x_{i+1} \cdots x_j| \neq |y_j \cdots y_{i+1}|$  then it contradicts to our conditions. Otherwise,  $||x_1 \cdots x_{i-1} x_j \cdots x_n| - |y_n \cdots y_j y_{i-1} \cdots y_1|| \geq |V|$  and

$$X_0 \xrightarrow{*}_G x_1 \cdots x_{i-1} x_j \cdots x_n X_n y_n \cdots y_j y_{i-1} \cdots y_1$$

also holds. Following this treatment, in finite steps we can reach  $X_0 \xrightarrow{*}_G a_1 \cdots a_k X b_k \cdots b_1, a_1, \dots, a_k, b_1, \dots, b_k \in \Sigma \cup \{\lambda\}$  with  $||a_1 \cdots a_k| - |b_k \cdots b_1|| \geq |V|$  such that  $k < |V|$ , which is impossible. Therefore, for every  $X \in V, p, q \in \Sigma^*, S \xrightarrow{*}_G pXq$  implies  $||p| - |q|| < |V|$ .

Now, for every derivation step, we order two pip-line stores, called *left store* and *right store*. Either both of them is empty, or one of them is empty and the another one contains a non-empty terminal string of length less than  $|V|$ .

At the start, both stores are empty. This status remains until the applied derivation rules are of the form  $X \rightarrow aYa, X, Y \in V, a \in \Sigma \cup \{\lambda\}$ . If the applied derivation rule has the form  $X \rightarrow aY, X, Y \in V, a \in \Sigma$ , then there are two cases: if the left store is empty, then we drop the terminal letter  $a$  into the top of the right store; otherwise we delete the terminal letter contained at

the bottom of the left store. (In the second case, the bottom of the left store should contain the same terminal letter  $a$ . Otherwise the generated word will not be palindrome.) Similarly, if the applied derivation rule has the form  $X \rightarrow Yb, X, Y \in V, b \in \Sigma$ , then we have two cases: if the left store is empty, then we drop the terminal letter  $b$  into the top of the left store; otherwise we delete the terminal letter contained at the bottom of the right store. ((In the second case again, the bottom of the left store should contain the same terminal letter  $b$ . Otherwise the generated word will not be palindrome.)

If the applied derivation rule has the form  $X \rightarrow aYb, X, Y \in V, a, b \in \Sigma$ , then we have the following possibilities: If one of the stores is not empty, then our procedure works as in the previous cases (like, in order, applying a derivation rule  $X \rightarrow aZ, a \in \Sigma, X, Z \in V$ , and then a derivation rule  $Z \rightarrow Yb, b \in \Sigma, Z, Y \in V$ ); if both stores are empty then  $a = b$  should hold. (Otherwise the generated string will not be palindrome.) After applying the considered derivation rule  $X \rightarrow aYb, X, Y \in V, a, b \in \Sigma$ , the contents of the stores remain the same.

We will construct our grammar such that a derivation rule of the form  $X \rightarrow a, a \in \Sigma \cup \{\lambda\}, X \in V$  can be applied only if either one of the stores contain the letter  $a$  or both stores are empty.

In addition, if both stores are empty, and  $X \xrightarrow[G]{*} w$  may hold for the non-terminal  $X$  contained on the left-hand side of the applied derivation rule, then  $w$  should be a palindrome. In addition, if  $|w| < |V|$ , then either  $w = b$  with  $b \in \Sigma \cup \{\lambda\}$ , or  $w = c_1 \cdots c_t d c_t \cdots c_1$  for some  $c_1, \dots, c_t \in \Sigma, d \in \Sigma \cup \{\lambda\}, 1 \leq t < |V|$ . For the second case, we assume the existence of some derivation rules of the form  $X \rightarrow c_1 Z_1 c_1, Z_1 \rightarrow c_2 Z_2 c_2, \dots, Z_{t-1} \rightarrow c_t Z_t c_t, Z_t \rightarrow d, Z_1, \dots, Z_t \in V$ .

Having this properties, formally we define the following derivation rules, where the (new) nonterminals are supplied by pile-line stores discussed previously.

Let  $\bar{V} = \{X \in V \mid X \xrightarrow[G]{*} w, w \in \Sigma^+, |w| < |V|\}$  and define, in order,  
 $V' = \{X_{\lambda, \lambda} \mid X \in V\} \cup \{X_{a_1 \dots a_k, \lambda} \mid X \in V, a_1, \dots, a_k \in \Sigma, k < |V|\}$   
 $\cup \{X_{\lambda, b_1 \dots b_k} \mid X \in V, b_1, \dots, b_k \in \Sigma, k < |V|\}$

and

$P' = \{X_{a_1 \dots a_k, \lambda} \rightarrow aY_{a_1 \dots a_k a, \lambda} a, X_{\lambda, a_1 \dots a_k} \rightarrow Y_{\lambda, a_1 \dots a_{k-1}}, X_{\lambda, \lambda} \rightarrow aY_{a, \lambda} a$   
 $\mid X \rightarrow Ya \in P, X, Y \in V, a_1, \dots, a_k, a \in \Sigma, k < |V|\} \cup$   
 $\{X_{a_1 \dots a_k, \lambda} \rightarrow Y_{a_1 \dots a_{k-1}, \lambda}, X_{\lambda, a_1 \dots a_k} \rightarrow aY_{\lambda, a_1 \dots a_k a} a, X_{\lambda, \lambda} \rightarrow aY_{\lambda, a} a$   
 $\mid X \rightarrow aY \in P, X, Y \in V, a_1, \dots, a_k, a \in \Sigma, k < |V|\} \cup$

$$\begin{aligned}
& \{X_{a_1 \dots a_k, \lambda} \rightarrow bY_{a_1 \dots a_{k-1} b, \lambda} b, X_{\lambda, a_1 \dots a_k} \rightarrow aY_{\lambda, a_1 \dots a_{k-1} a} a, X_{\lambda, \lambda} \rightarrow aY_{\lambda, \lambda} b \\
& \mid X \rightarrow aYb \in P, X, Y \in V, a_1, \dots, a_k, a, b \in \Sigma \cup \{\lambda\}\} \cup \\
& \{X_{a_1 \dots a_k, \lambda} \rightarrow Y_{a_1 \dots a_k, \lambda}, X_{\lambda, a_1 \dots a_k} \rightarrow Y_{\lambda, a_1 \dots a_k}, X_{\lambda, \lambda} \rightarrow Y_{\lambda, \lambda} \\
& \mid X \rightarrow Y \in P, X, Y \in V, a_1, \dots, a_k, \in \Sigma \cup \{\lambda\}\} \cup \{X_{a, \lambda} \rightarrow \lambda, X_{\lambda, a} \rightarrow \lambda, \\
& X_{\lambda, \lambda} \rightarrow a \mid X \rightarrow a \in P, X \in V, a \in \Sigma\} \cup \\
& \{X_{\lambda, \lambda} \rightarrow \lambda \mid X \rightarrow \lambda \in P\} \cup \{X_{\lambda, \lambda} \rightarrow c_1 Z_{1_{X\lambda, \lambda}} c_1, \\
& Z_{1_{X\lambda, \lambda}} \rightarrow c_2 Z_{2_{X\lambda, \lambda}} c_2, \dots, Z_{t-1_{X\lambda, \lambda}} \rightarrow c_t Z_{t_{X\lambda, \lambda}} c_t, Z_{t_{X\lambda, \lambda}} \rightarrow d \mid X \in \bar{V}, \\
& X \xrightarrow{G}^* c_1 \dots c_t d c_t \dots c_1, c_1, \dots, c_t \in \Sigma, d \in \Sigma \cup \{\lambda\}\}.
\end{aligned}$$

Thus we can receive that  $L(G) = L(G')$ , where  $G' = (V', \Sigma, S_{\lambda, \lambda}, P')$ .  $\square$

**Theorem 18** [5] *A context-free language  $L \subseteq \Sigma^*$  is palindromic if and only if it is a disjoint union of  $|V|$  number of languages of the form  $\{pap^R \mid p \in L_a\}$ , where the  $L_a$  ( $a \in \Sigma \cup \{\lambda\}$ ) are regular languages (uniquely determined by  $L$ ).*

*Proof:* Given an alphabet  $\Sigma$ , for every  $a \in \Sigma \cup \{\lambda\}$  consider a regular language  $L_a$ . It is clear that  $L = \bigcup_{a \in \Sigma \cup \{\lambda\}} \{pap^R : p \in L_a\}$  is palindromic and linear (and thus, it is also context-free). Conversely, consider a palindromic context-free language  $L$ . By Lemma 17, it can be generated by a grammar  $G = (V, \Sigma, S, P)$  having  $P \subseteq \{X \rightarrow aYa \mid X, Y \in V, a \in \Sigma\} \cup \{X \rightarrow a \mid X \in V, a \in \Sigma\} \cup \{X \rightarrow \lambda \mid X \in \Sigma\}$ . For every  $a \in \Sigma \cup \{\lambda\}$ , define the grammar  $G_a = (V, \Sigma, S, P_a)$  with  $P_a = P \setminus \{X \rightarrow b \mid b \in \Sigma \cup \{\lambda\}, b \neq a\}$ . Obviously,  $L(G) = \bigcup_{a \in \Sigma \cup \{\lambda\}} L(G_a)$ . Moreover, for every  $a, b \in \Sigma \cup \{\lambda\}$ ,  $L(G_a) \cap L(G_b) \neq \emptyset$  if and only if  $a = b$ . Therefore,  $L$  is a disjoint union of the languages  $L(G_a), a \in \Sigma \cup \{\lambda\}$ . By the construction of  $G_a, a \in \Sigma \cup \{\lambda\}$ , it is clear that  $G_{a, \ell} = (V, \Sigma, S, P_{a, \ell})$  with  $P_{a, \ell} = \{X \rightarrow Yb \mid X \rightarrow bYb \in P_a, X, Y \in V, a \in \Sigma\} \cup \{X \rightarrow b \mid X \rightarrow b \in P_a, X \in V, a \in \Sigma \cup \{\lambda\}\}$  is a regular language. Similarly,  $G_{a, r} = (V, \Sigma, S, P_{a, r})$  with  $P_{a, r} = \{X \rightarrow bY \mid X \rightarrow bYb \in P_a, X, Y \in V, a \in \Sigma\} \cup \{X \rightarrow b \mid X \rightarrow b \in P_a, X \in V, a \in \Sigma \cup \{\lambda\}\}$  is regular. Moreover,  $L_a = L(G_{a, \ell}) = L(G_{a, r})$ , and  $L = \bigcup_{a \in \Sigma \cup \{\lambda\}} \{pap^R : p \in L_a\}$ .  $\square$

Of course, every palindromic context-sensitive (phrase-structured) language has the form

$$L = \bigcup_{a \in \Sigma \cup \{\lambda\}} \{pap^R : p \in L(a)\},$$

where the  $L(a)$  ( $a \in \Sigma \cup \{\lambda\}$ ) are context-sensitive (phrase-structured) languages (uniquely determined by  $L$ ). Next we prove that unlike the regular

and context-free cases, the above languages  $L(a)$ ,  $a \in \Sigma \cup \{\lambda\}$  can be arbitrary context-sensitive (phrase-structured) languages.

**Theorem 19** *Given an alphabet  $\Sigma$ , for every  $a \in \Sigma \cup \{\lambda\}$  consider an arbitrary context-sensitive (phrase-structured) language  $L(a)$ . Then*

$$L = \bigcup_{a \in \Sigma \cup \{\lambda\}} \{pap^R : p \in L(a)\}$$

*is not only palindromic but context-sensitive (phrase-structured) as well.*

## References

- [1] Bar-Hillel, Y.; Perles, M.; Shamir, E.: On formal properties of simple phrase structure grammars. *Zeitschrift für Phonetik, Sprachwissenschaft, und Kommunikationsforschung*, **14** (1961), 143-177.
- [2] Fine, N. J.; Wilf, H. S.: Uniqueness theorems for periodic functions. *Proc. Am. Math. Soc.* **16** (1965), 109-114.
- [3] Ginsburg, S.; Spanier, E. H. : Bounded ALGOL-like languages. *Trans. Am. Math. Soc.*, **113** (1964), 333-368.
- [4] Ginsburg, S.; Rice, H. G. : Two families of languages related to ALGOL. *J. Assoc. Computing Machinery*, **9** (1962), 350-371.
- [5] Horváth, S.; Karhumäki, J.; Kleijn, J.: Results concerning palindromicity. (Mathematical aspects of informatics, Mägdesprung, 1986). *J. Inform. Process. Cybernet.* **23** (1987), no. 8-9, 441-451.
- [6] Ilie, L.: On a conjecture about slender context-free languages. *Theoret. Comput. Sci.*, **132** (1994), 427-434.
- [7] Latteux, M; Thierrin, G.: Semidiscrete context-free languages. *Internat. J. Comput. Math.* **14** (1983), 3-18.
- [8] Lyndon, R. C.; Schützenberger, M. P.: The equation  $a^m = b^n c^p$  in a free group. *Michigan Math. J.*, **9** (1962), 289-298.



- [9] Raz, D.: Length considerations in context-free languages. *Theoret. Comput. Sci.*, **183** (1997), 21–32.
- [10] Shyr, H. J.; Thierrin, G.: Disjunctive languages and codes. *In: Karpinński (ed.) : Proc. Conf. FCT'77*, **56** (1977), Springer-Verlag, 171–176.